REPORT DOCUMENTATION PAGE	Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
04-12-2014	Final Report	26-Aug-2011 - 25-Nov-2014

4. TITLE AND SUBTITLE
Final Report: A Unified Framework for Multi-level Processing of Complex Data

5a. CONTRACT NUMBER
W911NF-11-1-0426

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS
Charles K Chui

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of Missouri - St. Louis
Office of Research Administration
One University Boulevard
St. Louis, MO 63121 -4400

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
58539-CS.15

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
Complex data are lifted to a high-dimensional point-cloud for exploring data similarities, with each point representing an image thumb-nail, highlight of a medical record, spectral curve for every pixel of an HSI cube, etc. A weighted graph, with data similarities as weights, is constructed to connect the points of the point-cloud, and embedded to some binary tree by applying the shortest-path algorithm. The objective is to map the tree to the unit interval of the real-line, allowing us to extend the theory and methods from harmonic analysis to the study of

15. SUBJECT TERMS
Data geometry; Manifolds governed by point-clouds; Anisotropic transformation; Context-preservation extension; Tree-based data structure and mapping of big data; Harmonic analysis on point-clouds, graph and trees; Optimal filter bank; VM wavelets;

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Charles Chui
UU	UU	UU	UU		19b. TELEPHONE NUMBER 650-867-6510

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

Report Title

Final Report: A Unified Framework for Multi-level Processing of Complex Data

ABSTRACT

Complex data are lifted to a high-dimensional point-cloud for exploring data similarities, with each point representing an image thumb-nail, highlight of a medical record, spectral curve for every pixel of an HSI cube, etc. A weighted graph, with data similarities as weights, is constructed to connect the points of the point-cloud, and embedded to some binary tree by applying the shortest-path algorithm. The objective is to map the tree to the unit interval of the real-line, allowing us to extend the theory and methods from harmonic analysis to the study of functions on the given complex data. To build a unified framework for multi-level processing of the given complex data, spline and wavelet methods and algorithms have been developed with emphasis on real-time implementation. Toward the end of the funding period, an innovative theory, along with local methods, was developed for separating nonlinear and non-stationary signals from a blind source embedded with noise, via extraction of polynomial-like trends, point-set clustering, and estimation of instantaneous frequencies. This development has also been extended to the multivariate setting, including separation of image data.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received		Paper
08/13/2012	1.00	Bin Han, Charles K. Chui, Xiaosheng Zhuang. A Dual-chain Approach for Bottom-up Construction of Wavelet Filters with Any Integer Dilation, Applied Computational and Harmonic Analysis, (09 2012): 204. doi:
TOTAL:		**1**

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received		Paper
08/16/2013	4.00	J.Z. Wang, Charles Chui. Nonlinear methods for dimensionality reduction, Handbook of Geomathematics, W. Freeden, Z. Nashed, and Sonar T. (eds.), Springer, 2013, in press, (03 2013): 0. doi:
TOTAL:		**1**

Number of Papers published in non peer-reviewed journals:

(c) Presentations

1 Invited Speaker, two lectures: "Multi-level methods for image inpainting" and "Manifold approach to high-dimensional data processing," Università degli Studi Mediterranea di Reggio Calabria, Italy, July 4, 2012.

2. Invited Speaker, "Mathematical challenges," University of Messina, Sicily, Italy, July 10, 2012.

3. Invited Speaker, ``Multi-level methods for image inpainting," Università degli Studi Mediterranea di Reggio Calabria, Italy, July 4, 2012.

4. Invited Speaker, ``Manifold approach to high-dimensional data processing," Università degli Studi Mediterranea di Reggio Calabria, Italy, July 4, 2012.

5. Talk, ``Approximation on Unknown Manifolds defined by High-Dimensional Unstructured Data," Washington University, St. Louis, Nov. 2, 2012.

6. Colloquium talk, ``Mathematics for Data Recovery and Repair," UMSL, Nov. 14, 2012.

Number of Presentations: 6.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

TOTAL:

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

Received		Paper
08/08/2013	8.00	Charles Chui, H.N. Mhaskar. Smooth function extension based on high dimensional unstructured data, Mathematics of Computation, accepted and in press (09 2011)
08/12/2012	2.00	Charles K. Chui, H.N. Mhaskar. Smooth function extension based on high dimensional unstructured data, Mathematics of Computation, submitted 2012 (02 2012)
TOTAL:		**2**

Number of Manuscripts:

Books

Received	Book
TOTAL:	

Received	Book Chapter
TOTAL:	

Patents Submitted

Patents Awarded

Awards

1. Organizers: Akram Aldroubi, Charles Chui, and Alex Powell, "Fifth International Conference on Computational Harmonic Analysis", May 19-23, 2014, Vanderbilt University, Nashville, Tennessee.

2. Mathematics Conference Honoree, "International Conference on Interactions between \break Wavelets and Splines," in honor of the Birthday of Professor Charles K. Chui, May 16--19, 2005, University of Georgia, Athens, GA, with sponsorship and grants from the National Science Fundation (NSF) and the U.S. Army Research Office (ARO)

3. Shanks Endowed Lecture, Vanderbilt University, Nashville, May 18, 2011.

Graduate Students

NAME	PERCENT SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Post Doctorates

NAME	PERCENT SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Faculty Supported

NAME	PERCENT SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Under Graduate students supported

NAME	PERCENT SUPPORTED
FTE Equivalent:	
Total Number:	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: 0.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:...... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):...... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:...... 0.00

Names of Personnel receiving masters degrees

NAME

Total Number:

Names of personnel receiving PHDs

NAME

Total Number:

Names of other research staff

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Sub Contractors (DD882)

Inventions (DD882)

Scientific Progress

1. Theory and methods

Complex data are lifted to some high-dimensional point-cloud for exploring data similarities and data geometry. Examples include point-clouds of digital images, medical records, and hyperspectral image (HSI) cubes, where each point in the point-cloud could be an image thumb-nail, highlight of a medical record, and the spectral curve for each pixel of an HSI cube, respectively. There were two concurrent phases of our research program on this approach, with the first being theoretical development, and the second being development of innovative methods; but both toward the same goal of extending any desired objective function defined only on a subset (called training set) of the given point-cloud to the entire manifold, and then constructing an effective function representation of the extended objective function for mathematical analysis. Therefore, a unified framework can be formulated for multi-level processing of (an arbitrary function on) the given complex data. In our paper published in the American Mathematical Society journal, “Mathematics of Computation” in 2014, we have developed a complete theory for extending a desired target function on the training set to the (unknown) manifold which contains the point-cloud, such that the order of approximation is optimal based on certain smoothness function class, should the manifold be known. This theoretical extension is a two-stage process: first by constructing a data-driven optimal-order (heat) polynomial approximation, and then by blending the approximant with an interpolation operator. In our paper entitled, “Representation of functions on big data: graphs and trees”, accepted for publication by the Elsevier journal, “Applied and Computational Harmonic Analysis” and has been available on line since July 1, 2014, we use data similarities (obtained from a diffusion process, via our “anisotropic transform”) as weights to construct a weighted graph that connects the points of the point-cloud. We then embed this graph to a binary tree by applying the shortest-path algorithm, in order to construct a reversible transformation to map the graph to the unit interval of the real-line. This allows us to apply the theory and powerful methods from approximation theory and harmonic analysis to represent the target function on the given complex data.

2. Algorithms and computational schemes

When the training data-set is well chosen, the theory developed on function extension based on high-dimensional unstructured data, as mentioned above, can be applied. However, since training data selection is highly experimental, we also focused on computational methods, algorithm development, construction of optimal filters, and case studies. To better understand this popular new research direction, and particularly our own problem area, we have done an exhaustive literature search. Based on the searched results and on our own research findings, we have written two comprehensive tutorial papers, published as Springer Handbook chapters: one in 2012 and the other in 2014, on “feature extractions” and “nonlinear methods for data dimensionality reduction”. In the course of algorithm and computational scheme development, we also found it most productive, and hopefully with broader significant positive impact to both scientific and educational advancement, by writing a basic applied mathematics textbook with emphasis on spectral and Fourier methods, dimensionality reduction, data compression, wavelet analysis, and various applications. This book was published by Atlantis Press together with Springer in 2013. On computational and algorithmic development, we have completed three research papers: the first, being “A dual-chain approach for bottom-up construction of wavelet filters with arbitrary integer dilation”; the second, being “Multi-rate systems with shortest spline-wavelet filters”; and the third, being “Real-time dynamics acquisition from irregular samples with application to anesthesia evaluation”. The first paper was published in the journal, “Applied and Computational Harmonic Analysis” in 2012; the second was submitted to the same journal for publication earlier this year; and the third was submitted to the World Scientific journal, “Analysis and Applications” in July, 2014. In the second paper, we have constructed filter banks with arbitrarily number of sub-bands and any desirable order of vanishing moments, by deriving effective recursive formulas for computing the shortest filters. In the third paper mentioned above, we have constructed optimal-order interpolating local spline basis functions (of arbitrary spline order on irregular knot sequences) for real-time implementation, both on bounded and half-infinite time intervals, and have also introduced the notion and derived explicit formulations of “vanishing moment (VM)” wavelets of spline functions of any desired order and on arbitrary irregular knots, that have minimum support and maximum order of vanishing moments. The VM wavelets were also applied in the same paper to compute the synchrosqueezing transform (SST) in real-time and without the need of computing the derivative of the continuous wavelet transform. While the interpolating local spline basis functions are used to produce a continuous-time signal from the irregular samples, the SST of this continuous-time signal is used as the reference frequency for estimating the instantaneous frequencies, yielding the dynamics of the time series. We have applied this algorithm and real-time computational scheme to anesthesia evaluation from EEG data successfully, as analyzed in the same paper.

3. Application to signal and image separation from a blind source

Motivated by our success in real-time computation of the SST in the third paper discussed above, we have spent a great effort, since the spring of 2014, to develop a new approach to significantly improve the state-of-the-art theory and methods for signal (component) separation from a blind source. Let us first briefly discuss the background of this problem. Based on the continuous wavelet transform (CWT), the notion of synchrosqueezing transform (SST), introduced by Daubechies Lu, and Wu (DLW) in their 2011 paper, published in the journal, “Applied and Computational Harmonic Analysis”, provides a mostly dependable reference frequency for the estimation of all the instantaneous frequencies (IF’s) of a given (blind source) signal. The motivation of the original DLW paper was to give a mathematically sound alternative of the popular “empirical mode decomposition (EMD)” scheme, proposed by Huang et al, for decomposing a nonlinear and non-stationary signal into a hierarchy of intrinsic mode functions (IMF’s) and applying the Hilbert transform to extend each IMF to an amplitude-frequency

modulated signal in order to determine the IF of each IMF component of the EMD hierarchy. In our paper, “Signal decomposition and analysis via extraction of frequencies”, submitted to the same journal in the past summer, we introduce an innovative method to achieve a more ambitious goal than the SST approach, first by extracting the polynomial-like trend from the source signal and computing the exact number of signal components, then giving better estimates of the IF of each signal component, and finally separating the possibly non-stationary signal components from the source signal. Hence, our method avoids the need of guessing the number of IF’s for the SST approach. Furthermore, our computational scheme can be realized in near real-time, and our mathematical theory has direct extension to the multivariate setting. One main advantage of the SST approach is that reference the IF's so extracted is assured to be nonnegative. On the other hand, for the EMD approach, while the number K of IF's is already governed by the number of IMF's, it is unfortunate that the IF's of IMF's are sometimes negative. Other limitations of the EMD scheme include the need of adapting both the sifting process (for computing the IMF's) and the Hilbert transform (for the formulation the IF's from the analytic extension) to bounded and half--infinite time intervals. In our paper, “Signal analysis via instantaneous frequency estimation of signal components” under preparation, we have eliminated the above-mentioned limitations of these two approaches, by introducing a hybrid EMD-SST scheme, and significantly improved the quality and computational complexity. More precisely, using the EMD eliminates the need of guessing the number of IF’s (as required by the SST approach) and using the SST for the IF of each IMF is assured to be nonnegative. Furthermore, by applying the interpolating local spline basis functions and VM wavelets in our paper, “Real-time dynamics acquisition from irregular samples with application to anesthesia evaluation” discussed above, we eliminate the boundary artifacts introduced by artificial adaption of the sifting process of EMD at the boundary and the errors of analytic extension when the Hilbert transform is not taken on the entire real axis.

Research papers

1. B. Han, C.K. Chui, and X.S. Zhaung, A dual-chain approach for bottom-up construction of wavelet filters with arbitrary integer dilation, Appl. and Comp. Harm. Anal. 33 (2012), 204–225.
2. C.K. Chui and H. N. Mhaskar, Smooth function extension based on high dimensional unstructured data, Math. Comp. 83 (2014), 2865--2891.
3. C.K. Chui, F. Filber, and H.N. Mhaskar, Representation of functions on big data: graphs and trees, Appl. and Comp. Harm. Anal. Accepted. Available on line, July 1, 2014, under “Most downloaded papers”.
4. C.K. Chui, J.M. De Villiers, and X.S. Zhuang, Multi-rate systems with shortest spline-wavelet filters, submitted to Appl. and Comp. Harm. Anal.
5. C.K. Chui and H.N. Mhaskar, Signal decomposition and analysis via extraction of frequencies, submitted to Appl. and Comp. Harmonic Anal.
6. C. K. Chui, Y. Lin, and H.T. Wu, Real-time dynamics acquisition from irregular samples with application to anesthesia evaluation, submitted to Analysis and Applications.
7. C.K. Chui and M. van der Walt, Signal analysis via instantaneous frequency estimation of signal components, to be submitted to International J. on Geomathematics.

Tutorial papers

1. C. K. Chui and J.Z. Wang, Dimensionality Reduction of Hyperspectral Imagery Data for Feature Classification, in “Handbook of Geomathematics”, Chapter 47 (44 pages), Springer, 2010.
2. C. K. Chui and J. Z. Wang, Nonlinear methods for dimensionality reduction, in “Handbook of Geomathematics II”, Springer, In press.

Books

1. C.K. Chui and Q.T. Jiang, “Applied Mathematics: Data Compression, Spectral methods, Fourier Analysis, Wavelets, and Applications”, published by Atlantis Press/Springer, 555 pages + xxi , September, 2013 (ISBN 978-94-6239-008-9).

Technology Transfer